

SEQUENCE LISTING

(1) GENERAL INFORMATION

- (i) APPLICANT: Tartaglia, Louis A.
Weng, Xun
- (ii) TITLE OF THE INVENTION: NUCLEIC ACID MOLECULES ENCODING
GLUTEX AND USES THEREOF
- (iii) NUMBER OF SEQUENCES: 10
- (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Fish & Richardson P.C.
(B) STREET: 225 Franklin Street
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110-2804
- (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Diskette
(B) COMPUTER: IBM Compatible
(C) OPERATING SYSTEM: Windows95
(D) SOFTWARE: FastSEQ for Windows Version 2.0
- (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: 09/299,349
(B) FILING DATE: 26-APR-1999
- (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 09/031,392
(B) FILING DATE: 26-FEB-1998
- (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Meiklejohn, Ph.D., Anita L.
(B) REGISTRATION NUMBER: 35,283
(C) REFERENCE/DOCKET NUMBER: 07334/072002
- (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 617/542-5070
(B) TELEFAX: 617/542-8906
(C) TELEX: 200154

(2) INFORMATION FOR SEQ ID NO:1:

- (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2343 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: cDNA
- (ix) FEATURE:
(A) NAME/KEY: Coding Sequence
(B) LOCATION: 73...1761

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

TCGACCCACG CGTCCGGCCT TGGCAGAGTC TGGGGTCCCT GGA CTGAGCC ATCAGCTGGG 60
TCACTGAGAC CC ATG GCA AGG AAA CAA AAT AGG AAT TCC AAG GAA CTG GGC 111
Met Ala Arg Lys Gln Asn Arg Asn Ser Lys Glu Leu Gly
1 5 10

CTA GTT CCC CTC ACA GAT GAC ACC AGC CAC GCC GGG CCT CCA GGG CCA 159
Leu Val Pro Leu Thr Asp Asp Thr Ser His Ala Gly Pro Pro Gly Pro
15 20 25

GGG AGG GCA CTG CTG GAG TGT GAC CAC CTG AGG AGT GGG GTG CCA GGT 207
Gly Arg Ala Leu Leu Glu Cys Asp His Leu Arg Ser Gly Val Pro Gly
30 35 40 45

GGA AGG AGA AGA AAG GAC TGG TCC TGC TCG CTC CTC GTG GCC TCC CTC 255
Gly Arg Arg Arg Lys Asp Trp Ser Cys Ser Leu Leu Val Ala Ser Leu
50 55 60

GCG GGC GCC TTC GGC TCC TCC TTC CTC TAC GGC TAC AAC CTG TCG GTG 303
Ala Gly Ala Phe Gly Ser Ser Phe Leu Tyr Gly Tyr Asn Leu Ser Val
65 70 75

GTG AAT GCC CCC ACC CCG TAC ATC AAG GCC TTT TAC AAT GAG TCA TGG 351
Val Asn Ala Pro Thr Pro Tyr Ile Lys Ala Phe Tyr Asn Glu Ser Trp
80 85 90

GAA AGA AGG CAT GGA CGT CCA ATA GAC CCA GAC ACT CTG ACT CTG CTC 399
Glu Arg Arg His Gly Arg Pro Ile Asp Pro Asp Thr Leu Thr Leu Leu
95 100 105

TGG TCT GTG ACT GTG TCC ATA TTC GCC ATC GGT GGA CTT GTG GGG ACG 447
Trp Ser Val Thr Val Ser Ile Phe Ala Ile Gly Gly Leu Val Gly Thr
110 115 120 125

TTA ATT GTG AAG ATG ATT GGA AAG GTT CTT GGG AGG AAG CAC ACT TTG 495
Leu Ile Val Lys Met Ile Gly Lys Val Leu Gly Arg Lys His Thr Leu
130 135 140

CTG GCC AAT AAT GGG TTT GCA ATT TCT GCT GCA TTG CTG ATG GCC TGC 543
Leu Ala Asn Asn Gly Phe Ala Ile Ser Ala Ala Leu Leu Met Ala Cys
145 150 155

TCG CTC CAG GCA GGA GCC TTT GAA ATG CTC ATT GTG GGA CGC TTC ATC 591
Ser Leu Gln Ala Gly Ala Phe Glu Met Leu Ile Val Gly Arg Phe Ile
160 165 170

ATG GGC ATA GAT GGA GGC GTC GCC CTC AGT GTG CTC CCC ATG TAC CTC 639
Met Gly Ile Asp Gly Gly Val Ala Leu Ser Val Leu Pro Met Tyr Leu
175 180 185

AGT GAG ATC TCA CCC AAG GAG ATC CGT GGC TCT CTG GGG CAG GTG ACT 687
Ser Glu Ile Ser Pro Lys Glu Ile Arg Gly Ser Leu Gly Gln Val Thr
190 195 200 205

GCC ATC TTT ATC TGC ATT GGC GTG TTC ACT GGG CAG CTT CTG GGC CTG 735
Ala Ile Phe Ile Cys Ile Gly Val Phe Thr Gly Gln Leu Leu Gly Leu
210 215 220

CCC GAG CTG CTG GGA AAG GAG AGT ACC TGG CCA TAC CTG TTT GGA GTG 783
Pro Glu Leu Leu Gly Lys Glu Ser Thr Trp Pro Tyr Leu Phe Gly Val
225 230 235

ATT GTG GTC CCT GCC GTT GTC CAG CTG CTG AGC CTT CCC TTT CTC CCG	831
Ile Val Val Pro Ala Val Val Gln Leu Leu Ser Leu Pro Phe Leu Pro	
240 245 250	
GAC AGC CCA CGC TAC CTG CTC TTG GAG AAG CAC AAC GAG GCA AGA GCT	879
Asp Ser Pro Arg Tyr Leu Leu Leu Glu Lys His Asn Glu Ala Arg Ala	
255 260 265	
GTG AAA GCC TTC CAA ACG TTC TTG GGT AAA GCA GAC GTT TCC CAA GAG	927
Val Lys Ala Phe Gln Thr Phe Leu Gly Lys Ala Asp Val Ser Gln Glu	
270 275 280 285	
GTA GAG GAG GTC CTG GCT GAG AGC CAC GTG CAG AGG AGC ATC CGC CTG	975
Val Glu Glu Val Leu Ala Glu Ser His Val Gln Arg Ser Ile Arg Leu	
290 295 300	
GTG TCC GTG CTG GAG CTG CTG AGA GCT CCC TAC GTC CGC TGG CAG GTG	1023
Val Ser Val Leu Glu Leu Leu Arg Ala Pro Tyr Val Arg Trp Gln Val	
305 310 315	
GTC ACC GTG ATT GTC ACC ATG GCC TGC TAC CAG CTC TGT GGC CTC AAT	1071
Val Thr Val Ile Val Thr Met Ala Cys Tyr Gln Leu Cys Gly Leu Asn	
320 325 330	
GCA ATT TGG TTC TAT ACC AAC AGC ATC TTT GGA AAA GCT GGG ATC CCT	1119
Ala Ile Trp Phe Tyr Thr Asn Ser Ile Phe Gly Lys Ala Gly Ile Pro	
335 340 345	
CCG GCA AAG ATC CCA TAC GTC ACC TTG AGT ACA GGG GGC ATC GAG ACT	1167
Pro Ala Lys Ile Pro Tyr Val Thr Leu Ser Thr Gly Gly Ile Glu Thr	
350 355 360 365	
TTG GCT GCC GTC TTC TCT GGT TTG GTC ATT GAG CAC CTG GGA CGG AGA	1215
Leu Ala Ala Val Phe Ser Gly Leu Val Ile Glu His Leu Gly Arg Arg	
370 375 380	
CCC CTC CTC ATT GGT GGC TTT GGG CTC ATG GGC CTC TTC TTT GGG ACC	1263
Pro Leu Leu Ile Gly Gly Phe Gly Leu Met Gly Leu Phe Phe Gly Thr	
385 390 395	
CTC ACC ATC ACG CTG ACC CTG CAG GAC CAC GCC CCC TGG GTC CCC TAC	1311
Leu Thr Ile Thr Leu Thr Leu Gln Asp His Ala Pro Trp Val Pro Tyr	
400 405 410	
CTG AGT ATC GTG GGC ATT CTG GCC ATC ATC GCC TCT TTC TGC AGT GGG	1359
Leu Ser Ile Val Gly Ile Leu Ala Ile Ile Ala Ser Phe Cys Ser Gly	
415 420 425	
CCA GGT GGC ATC CCG TTC ATC TTG ACT GGT GAG TTC TTC CAG CAA TCT	1407
Pro Gly Gly Ile Pro Phe Ile Leu Thr Gly Glu Phe Phe Gln Gln Ser	
430 435 440 445	
CAG CGG CCG GCT GCC TTC ATC ATT GCA GGC ACC GTC AAC TGG CTC TCC	1455
Gln Arg Pro Ala Ala Phe Ile Ile Ala Gly Thr Val Asn Trp Leu Ser	
450 455 460	
AAC TTT GCT GTT GGG CTC CTC TTC CCA TTC ATT CAG AAA AGT CTG GAC	1503
Asn Phe Ala Val Gly Leu Leu Phe Pro Phe Ile Gln Lys Ser Leu Asp	
465 470 475	

ACC TAC TGT TTC CTA GTC TTT GCT ACA ATT TGT ATC ACA GGT GCT ATC	1551
Thr Tyr Cys Phe Leu Val Phe Ala Thr Ile Cys Ile Thr Gly Ala Ile	
480 485 490	
TAC CTG TAT TTT GTG CTG CCT GAG ACC AAA AAC AGA ACC TAT GCA GAA	1599
Tyr Leu Tyr Tyr Phe Val Leu Pro Glu Thr Lys Asn Arg Thr Tyr Ala Glu	
495 500 505	
ATC AGC CAG GCA TTT TCC AAA AGG AAC AAA GCA TAC CCA CCA GAA GAG	1647
Ile Ser Gln Ala Phe Ser Lys Arg Asn Lys Ala Tyr Pro Pro Glu Glu	
510 515 520 525	
AAA ATC GAC TCA GCT GTC ACT GAT GCT CCT GCT TCT TCT CCT TTC ACT	1695
Lys Ile Asp Ser Ala Val Thr Asp Ala Pro Ala Ser Ser Pro Phe Thr	
530 535 540	
ACT CCG AAT ACA GCC TGG ATT CAA GCT GCC GCC ACC ACC ACC GCC ACC	1743
Thr Pro Asn Thr Ala Trp Ile Gln Ala Ala Ala Thr Thr Thr Ala Thr	
545 550 555	
AAA AAA GAA CAC CCA TTG TAAACGGTCA TGTGGTATTT CCTCAACCTG GAATGACC	1799
Lys Lys Glu His Pro Leu	
560	
TTCCCCTATC TTCTTCTCCT GGAGAACACC AAGTCATGAT GTCAGACAAG AGCTTGGATT	1859
TTGGAGACAT GGGTTTGAAT TCCAGTCATT CATTCTTTTA TTCAGCAAAT ATTTAACAAG	1919
TACTGACATG TCCCATATGT TGTTTTACCC ACTGGTTATA CAATGGGAGG GAGAGAGAGA	1979
GAGAGAGAGA GAGAGAGATG CTATTCTAAA AGCTTGAAGT CTAGGCTGTG CACGGTGGCT	2039
CACGCCTGTA ATCCCAGCAC TTTGGGAGGC CGAGGTGGGT GGATCGTGAG GTCAGGAGAT	2099
TGAGACCATC CTGGCTAACA TGGTGAAACT CCCTCTCTAC TAAAAATACA AAAAATTAGC	2159
TGAGCATGGT GGCGGGCGCC TGTAGTCCCA GCTACTTGGG AGGCTGAGGC AGGAGAATGG	2219
CGTGAACCCA GGAGGCGGAG CTTGCAGTGA GCCGAGATCA CACCACCACA CTCCAGCCTG	2279
GGTGACAGAG CCAGACTCCG TCTCAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAGGGCGG	2339
CCGC	2343

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 563 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met	Ala	Arg	Lys	Gln	Asn	Arg	Asn	Ser	Lys	Glu	Leu	Gly	Leu	Val	Pro
1				5					10					15	
Leu	Thr	Asp	Asp	Thr	Ser	His	Ala	Gly	Pro	Pro	Gly	Pro	Gly	Arg	Ala
		20						25					30		
Leu	Leu	Glu	Cys	Asp	His	Leu	Arg	Ser	Gly	Val	Pro	Gly	Gly	Arg	Arg
		35					40					45			
Arg	Lys	Asp	Trp	Ser	Cys	Ser	Leu	Leu	Val	Ala	Ser	Leu	Ala	Gly	Ala
	50				55				60						
Phe	Gly	Ser	Ser	Phe	Leu	Tyr	Gly	Tyr	Asn	Leu	Ser	Val	Val	Asn	Ala
65				70					75					80	
Pro	Thr	Pro	Tyr	Ile	Lys	Ala	Phe	Tyr	Asn	Glu	Ser	Trp	Glu	Arg	Arg
			85						90				95		
His	Gly	Arg	Pro	Ile	Asp	Pro	Asp	Thr	Leu	Thr	Leu	Leu	Trp	Ser	Val
			100					105					110		

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Met	Gly	Phe	Ser	Lys	Leu	Gly	Lys	Ser	Phe	Glu	Met	Leu	Ile	Leu	Gly
1				5				10					15		
Arg	Phe	Ile	Ile	Gly	Val	Tyr	Cys	Gly	Leu	Thr	Thr	Gly	Phe	Val	Pro
		20					25						30		
Met	Tyr	Val	Gly	Glu	Val	Ser	Pro	Thr	Glu	Leu	Arg	Gly	Ala	Leu	Gly
		35					40					45			
Thr	Leu	His	Gln	Leu	Gly	Ile	Val	Val	Gly	Ile	Leu	Ile	Ala	Gln	Val
	50					55				60					
Phe	Gly	Leu	Asp	Ser	Ile	Met	Gly	Asn	Gln	Glu	Leu	Trp	Pro	Leu	Leu
65					70				75					80	
Leu	Ser	Val	Ile	Phe	Ile	Pro	Ala	Leu	Leu	Gln	Cys	Ile	Leu	Leu	Pro
			85					90					95		
Phe	Cys	Pro	Glu	Ser	Pro	Arg	Phe	Leu	Leu	Ile	Asn	Arg	Asn	Glu	Glu
		100						105					110		
Asn	Arg	Ala	Lys	Ser	Val	Leu	Lys	Lys	Leu	Arg	Gly	Thr	Ala	Asp	Val
	115						120					125			
Thr	Arg	Asp	Leu	Gln	Glu	Met	Lys	Glu	Glu	Ser	Arg	Gln	Met	Met	Arg
	130					135				140					
Glu	Lys	Lys	Val	Thr	Ile	Leu	Glu	Leu	Phe	Arg	Ser	Ala	Ala	Tyr	Arg
145					150				155					160	
Gln	Pro	Ile	Leu	Ile	Ala	Val	Val	Leu	Gln	Leu	Ser	Gln	Gln	Leu	Ser
			165					170						175	
Gly	Ile	Asn	Ala	Val	Phe	Tyr	Tyr	Ser	Thr	Ser	Ile	Phe	Glu	Lys	Ala
		180						185					190		
Gly	Val	Gln	Pro	Val	Tyr	Ala	Thr	Ile	Gly	Ser	Gly	Ile	Val	Asn	
	195					200					205				
Thr	Ala	Phe	Thr	Val	Val	Ser	Leu	Phe	Val	Val	Glu	Arg	Ala	Gly	Arg
	210					215					220				
Arg	Thr	Leu	His	Leu	Ile	Gly	Leu	Ala	Gly	Met	Ala	Gly	Cys	Ala	Val
225					230					235				240	
Leu	Met	Thr	Ile	Ala	Leu	Ala	Leu	Leu	Glu	Gln	Leu	Pro	Trp	Met	Ser
			245						250					255	
Tyr	Leu	Ser	Ile	Val	Ala	Ile	Phe	Gly	Phe	Val	Ala	Phe	Phe	Glu	Val
		260						265					270		
Gly	Pro	Gly	Pro	Ile	Pro	Trp	Phe	Ile	Val	Ala	Glu	Leu	Phe	Ser	Gln
		275					280					285			
Gly	Pro	Arg	Pro	Ala	Ala	Ile	Ala	Val	Ala	Gly	Phe	Ser	Asn	Trp	Thr
	290					295					300				
Ser	Asn	Phe	Ile	Val	Gly	Met	Cys	Phe	Gln	Tyr	Val	Glu	Gln	Leu	Cys
305					310					315				320	
Gly	Pro	Tyr	Val	Phe	Ile	Ile	Phe	Thr	Val	Leu	Leu	Val	Leu	Phe	Phe
			325					330						335	
Ile	Phe	Thr	Tyr	Phe	Lys	Val	Pro	Glu	Thr	Lys	Gly	Arg	Thr	Phe	Asp
		340						345					350		
Glu	Ile	Ala	Ser	Gly	Phe	Arg	Gln	Gly	Gly	Ala	Ser	Gln	Ser	Asp	Lys
		355					360					365			
Thr	Pro	Glu	Glu	Leu	Phe	His	Pro	Leu	Gly	Ala	Asp	Ser	Gln	Val	
	370					375					380				

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 534 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Asp	Gly	Lys	Ser	Lys	Met	Gln	Ala	Glu	Lys	His	Leu	Thr	Gly	Thr
1				5					10					15	
Leu	Val	Leu	Ser	Val	Phe	Thr	Ala	Val	Leu	Gly	Phe	Phe	Gln	Tyr	Gly
			20					25					30		
Tyr	Ser	Leu	Gly	Val	Ile	Asn	Ala	Pro	Gln	Lys	Val	Ile	Glu	Ala	His
		35					40					45			
Tyr	Gly	Arg	Met	Leu	Gly	Ala	Ile	Pro	Met	Val	Arg	His	Ala	Thr	Asn
	50					55					60				
Thr	Ser	Arg	Asp	Asn	Ala	Thr	Ile	Thr	Val	Thr	Ile	Pro	Gly	Thr	Glu
65				70						75					80
Ala	Trp	Gly	Ser	Ser	Glu	Gly	Thr	Leu	Ala	Pro	Ser	Ala	Gly	Phe	Glu
				85					90					95	
Asp	Pro	Thr	Val	Ser	Pro	His	Ile	Leu	Thr	Met	Tyr	Trp	Ser	Leu	Ser
			100					105					110		
Val	Ser	Met	Phe	Ala	Val	Gly	Gly	Met	Val	Ser	Ser	Phe	Thr	Val	Gly
		115					120					125			
Trp	Ile	Gly	Asp	Arg	Leu	Gly	Arg	Val	Lys	Ala	Met	Leu	Val	Val	Asn
130					135						140				
Val	Leu	Ser	Ile	Ala	Gly	Asn	Leu	Leu	Met	Gly	Leu	Ala	Lys	Met	Gly
145					150					155					160
Pro	Ser	His	Ile	Leu	Ile	Ile	Ala	Gly	Arg	Ala	Ile	Thr	Gly	Leu	Tyr
			165					170						175	
Cys	Gly	Leu	Ser	Ser	Gly	Leu	Val	Pro	Met	Tyr	Val	Ser	Glu	Val	Ser
			180					185					190		
Pro	Thr	Ala	Leu	Arg	Gly	Ala	Leu	Gly	Thr	Leu	His	Gln	Leu	Ala	Ile
		195					200					205			
Val	Thr	Gly	Ile	Leu	Ile	Ser	Gln	Val	Leu	Gly	Leu	Asp	Phe	Leu	Leu
210					215						220				
Gly	Asn	Asp	Glu	Leu	Trp	Pro	Leu	Leu	Leu	Gly	Leu	Ser	Gly	Val	Ala
225					230					235					240
Ala	Leu	Leu	Gln	Phe	Phe	Leu	Leu	Leu	Leu	Cys	Pro	Glu	Ser	Pro	Arg
			245					250						255	
Tyr	Leu	Tyr	Ile	Lys	Leu	Gly	Lys	Val	Glu	Glu	Ala	Lys	Lys	Ser	Leu
			260				265						270		
Lys	Arg	Leu	Arg	Gly	Asn	Cys	Asp	Pro	Met	Lys	Glu	Ile	Ala	Glu	Met
		275				280						285			
Glu	Lys	Glu	Lys	Gln	Glu	Ala	Ala	Ser	Glu	Lys	Arg	Val	Ser	Ile	Gly
	290					295					300				
Gln	Leu	Phe	Ser	Ser	Ser	Lys	Tyr	Arg	Gln	Ala	Val	Ile	Val	Ala	Leu
305					310					315					320
Met	Val	Gln	Ile	Ser	Gln	Gln	Phe	Ser	Gly	Ile	Asn	Ala	Ile	Phe	Tyr
			325						330					335	
Tyr	Ser	Thr	Asn	Ile	Phe	Gln	Arg	Ala	Gly	Val	Gly	Gln	Pro	Val	Tyr
		340					345						350		
Tyr	Ala	Thr	Ile	Gly	Val	Gly	Val	Val	Asn	Thr	Val	Phe	Thr	Val	Ile
		355					360					365			
Ser	Val	Phe	Leu	Val	Glu	Lys	Ala	Gly	Arg	Arg	Ser	Leu	Phe	Leu	Ala
	370					375					380				
Gly	Leu	Met	Gly	Met	Leu	Ile	Ser	Ala	Val	Ala	Met	Thr	Val	Gly	Leu
385					390					395					400
Val	Leu	Leu	Ser	Gln	Phe	Ala	Trp	Met	Ser	Tyr	Val	Ser	Met	Val	Ala
			405					410						415	
Ile	Phe	Leu	Phe	Val	Ile	Phe	Phe	Glu	Val	Gly	Pro	Gly	Pro	Ile	Pro
		420						425					430		
Trp	Phe	Ile	Val	Ala	Glu	Leu	Phe	Ser	Gln	Gly	Pro	Arg	Pro	Ala	Ala
	435					440						445			
Ile	Ala	Val	Ala	Gly	Phe	Cys	Asn	Trp	Ala	Cys	Asn	Phe	Ile	Val	Gly
	450					455					460				

Met	Cys	Phe	Gln	Tyr	Ile	Ala	Asp	Leu	Cys	Gly	Pro	Tyr	Val	Phe	Val
465					470					475					480
Val	Phe	Ala	Val	Leu	Leu	Leu	Val	Phe	Phe	Leu	Phe	Ala	Tyr	Leu	Lys
				485						490					495
Val	Pro	Glu	Thr	Lys	Gly	Lys	Ser	Phe	Glu	Glu	Ile	Ala	Ala	Ala	Phe
			500					505					510		
Arg	Arg	Lys	Lys	Leu	Pro	Ala	Lys	Ser	Met	Thr	Glu	Leu	Glu	Asp	Leu
		515					520					525			
Arg	Gly	Gly	Glu	Glu	Ala										
	530														

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 494 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met	Gly	Thr	Thr	Lys	Val	Thr	Thr	Pro	Leu	Ile	Phe	Ala	Ile	Ser	Ile
1				5					10					15	
Ala	Thr	Ile	Gly	Ser	Phe	Gln	Phe	Gly	Tyr	Asn	Thr	Gly	Val	Ile	Asn
			20					25					30		
Ala	Pro	Glu	Ala	Ile	Ile	Lys	Asp	Phe	Leu	Asn	Tyr	Thr	Leu	Glu	Glu
		35					40					45			
Arg	Ser	Glu	Thr	Pro	Pro	Ser	Ser	Val	Leu	Leu	Thr	Ser	Leu	Trp	Ser
	50					55					60				
Leu	Ser	Val	Ala	Ile	Phe	Ser	Val	Gly	Gly	Met	Ile	Gly	Ser	Phe	Ser
65				70					75						80
Val	Gly	Leu	Phe	Val	Asn	Arg	Phe	Gly	Arg	Arg	Asn	Ser	Met	Leu	Ile
			85					90					95		
Val	Asn	Leu	Leu	Ala	Ile	Ala	Gly	Gly	Cys	Leu	Met	Gly	Phe	Cys	Lys
			100				105					110			
Ile	Ala	Glu	Ser	Val	Glu	Met	Leu	Ile	Leu	Gly	Arg	Leu	Ile	Ile	Gly
	115					120					125				
Leu	Phe	Cys	Gly	Leu	Cys	Thr	Gly	Phe	Val	Pro	Met	Tyr	Ile	Gly	Glu
	130					135					140				
Ile	Ser	Pro	Thr	Ala	Leu	Arg	Gly	Ala	Phe	Gly	Thr	Leu	Asn	Gln	Leu
145				150					155						160
Gly	Ile	Val	Ile	Gly	Ile	Leu	Val	Ala	Gln	Ile	Phe	Gly	Leu	Lys	Val
			165					170					175		
Ile	Leu	Gly	Thr	Glu	Asp	Leu	Trp	Pro	Leu	Leu	Leu	Gly	Phe	Thr	Ile
		180					185					190			
Leu	Pro	Ala	Ile	Ile	Gln	Cys	Ala	Ala	Leu	Pro	Phe	Cys	Pro	Glu	Ser
		195					200					205			
Pro	Arg	Phe	Leu	Leu	Ile	Asn	Arg	Lys	Glu	Glu	Glu	Lys	Ala	Lys	Glu
	210					215					220				
Ile	Leu	Gln	Arg	Leu	Trp	Gly	Thr	Glu	Asp	Val	Ala	Gln	Asp	Ile	Gln
225				230					235						240
Glu	Met	Lys	Asp	Glu	Ser	Met	Arg	Met	Ser	Gln	Glu	Lys	Gln	Val	Thr
			245					250					255		
Val	Leu	Glu	Leu	Phe	Arg	Ala	Pro	Asn	Tyr	Arg	Gln	Pro	Ile	Ile	Ile
		260					265					270			
Ser	Ile	Met	Leu	Gln	Leu	Ser	Gln	Leu	Ser	Gly	Ile	Asn	Ala	Val	
		275					280				285				
Phe	Tyr	Tyr	Ser	Thr	Gly	Ile	Phe	Lys	Asp	Ala	Gly	Val	Gln	Glu	Pro
	290				295					300					
Val	Tyr	Ala	Thr	Ile	Gly	Ala	Gly	Val	Val	Asn	Thr	Ile	Phe	Thr	Val
305					310					315					320

Val	Ser	Val	Phe	Leu	Val	Glu	Arg	Ala	Gly	Arg	Arg	Thr	Leu	His	Leu
				325					330					335	
Ile	Gly	Leu	Gly	Gly	M t	Ala	Phe	Cys	Ser	Ile	Leu	Met	Thr	Ile	Ser
			340					345					350		
Leu	Leu	Leu	Lys	Asp	Asn	Tyr	Ser	Trp	Met	Ser	Phe	Ile	Cys	Ile	Gly
			355				360					365			
Ala	Ile	Leu	Val	Phe	Val	Ala	Phe	Phe	Glu	Ile	Gly	Pro	Gly	Pro	Ile
			370			375					380				
Pro	Trp	Phe	Ile	Val	Ala	Glu	Leu	Phe	Gly	Gln	Gly	Pro	Arg	Pro	Ala
	385				390				395						400
Ala	Met	Ala	Val	Ala	Gly	Cys	Ser	Asn	Trp	Thr	Ser	Asn	Phe	Leu	Val
				405				410						415	
Gly	Leu	Leu	Phe	Pro	Ser	Ala	Thr	Phe	Tyr	Leu	Gly	Ala	Tyr	Val	Phe
			420					425					430		
Ile	Val	Phe	Thr	Val	Phe	Leu	Val	Ile	Phe	Trp	Val	Phe	Thr	Phe	Phe
		435				440					445				
Lys	Val	Pro	Glu	Thr	Arg	Gly	Arg	Thr	Phe	Glu	Glu	Ile	Thr	Arg	Ala
	450					455				460					
Phe	Glu	Gly	Gln	Val	Gln	Thr	Gly	Thr	Arg	Gly	Glu	Lys	Gly	Pro	Ile
	465				470				475						480
Met	Glu	Met	Asn	Ser	Ile	Gln	Pro	Thr	Lys	Asp	Thr	Asn	Ala		
			485					490							

(2) INFORMATION FOR SEQ ID NO:6:

- (i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 509 amino acids
 (B) TYPE: amino acid
 (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met	Pro	Ser	Gly	Phe	Gln	Gln	Ile	Gly	Ser	Glu	Asp	Gly	Glu	Pro	Pro
1				5				10						15	
Gln	Gln	Arg	Val	Thr	Gly	Thr	Leu	Val	Leu	Ala	Val	Phe	Ser	Ala	Val
			20					25					30		
Leu	Gly	Ser	Leu	Gln	Phe	Gly	Tyr	Asn	Ile	Gly	Val	Ile	Asn	Ala	Pro
		35				40						45			
Gln	Lys	Val	Ile	Glu	Gln	Ser	Tyr	Asn	Glu	Thr	Trp	Leu	Gly	Arg	Gln
	50				55				60						
Gly	Pro	Glu	Gly	Pro	Ser	Ser	Ile	Pro	Pro	Gly	Thr	Leu	Thr	Thr	Leu
	65				70				75					80	
Trp	Ala	Leu	Ser	Val	Ala	Ile	Phe	Ser	Val	Gly	Gly	Met	Ile	Ser	Ser
			85					90						95	
Phe	Leu	Ile	Gly	Ile	Ile	Ser	Gln	Trp	Leu	Gly	Arg	Lys	Arg	Ala	Met
			100				105						110		
Leu	Val	Asn	Asn	Val	Leu	Ala	Val	Leu	Gly	Gly	Ser	Leu	Met	Gly	Leu
		115				120						125			
Ala	Asn	Ala	Ala	Ala	Ser	Tyr	Glu	Met	Leu	Ile	Leu	Gly	Arg	Phe	Leu
	130				135						140				
Ile	Gly	Ala	Tyr	Ser	Gly	Leu	Thr	Ser	Gly	Leu	Val	Pro	Met	Tyr	Val
	145				150				155					160	
Gly	Glu	Ile	Ala	Pro	Thr	His	Leu	Arg	Gly	Ala	Leu	Gly	Thr	Leu	Asn
			165				170							175	
Gln	Leu	Ala	Ile	Val	Ile	Gly	Ile	Leu	Ile	Ala	Gln	Val	Leu	Gly	Leu
			180				185						190		
Glu	Ser	Leu	Leu	Gly	Thr	Ala	Ser	Leu	Trp	Pro	Leu	Leu	Leu	Gly	Leu
	195					200					205				
Thr	Val	Leu	Pro	Ala	Leu	Leu	Gln	Leu	Val	Leu	Leu	Pro	Phe	Cys	Pro
	210					215					220				

09081947 "1018001

Glu	Ser	Pro	Arg	Tyr	Leu	Tyr	Ile	Ile	Gln	Asn	Leu	Glu	Gly	Pro	Ala
225					230					235					240
Arg	Lys	Ser	Leu	Lys	Arg	Leu	Thr	Gly	Trp	Ala	Asp	Val	Ser	Gly	Val
				245					250					255	
Leu	Ala	Glu	Leu	Lys	Asp	Glu	Lys	Arg	Lys	Leu	Glu	Arg	Glu	Arg	Pro
			260					265					270		
Leu	Ser	Leu	Leu	Gln	Leu	Leu	Gly	Ser	Arg	Thr	His	Arg	Gln	Pro	Leu
		275					280					285			
Ile	Ile	Ala	Val	Val	Leu	Gln	Leu	Ser	Gln	Gln	Leu	Ser	Gly	Ile	Asn
	290					295					300				
Ala	Val	Phe	Tyr	Tyr	Ser	Thr	Ser	Ile	Phe	Glu	Thr	Ala	Gly	Val	Gly
305					310					315					320
Gln	Pro	Ala	Tyr	Ala	Thr	Ile	Gly	Ala	Gly	Val	Val	Asn	Thr	Val	Phe
				325					330					335	
Thr	Leu	Val	Ser	Val	Leu	Leu	Val	Glu	Arg	Ala	Gly	Arg	Arg	Thr	Leu
			340					345					350		
His	Leu	Leu	Gly	Leu	Ala	Gly	Met	Cys	Gly	Cys	Ala	Ile	Leu	Met	Thr
		355				360						365			
Val	Ala	Leu	Leu	Leu	Leu	Glu	Arg	Val	Pro	Ala	Met	Ser	Tyr	Val	Ser
	370					375					380				
Ile	Val	Ala	Ile	Phe	Gly	Phe	Val	Ala	Phe	Phe	Glu	Ile	Gly	Pro	Gly
385					390					395					400
Pro	Ile	Pro	Trp	Phe	Ile	Val	Ala	Glu	Leu	Phe	Ser	Gln	Gly	Pro	Arg
				405				410						415	
Pro	Ala	Ala	Met	Ala	Val	Ala	Gly	Phe	Ser	Asn	Trp	Thr	Ser	Asn	Phe
			420					425					430		
Ile	Ile	Gly	Met	Gly	Phe	Gln	Tyr	Val	Ala	Glu	Ala	Met	Gly	Pro	Tyr
		435				440						445			
Val	Phe	Leu	Leu	Phe	Ala	Val	Leu	Leu	Leu	Gly	Phe	Phe	Ile	Phe	Thr
	450				455					460					
Phe	Leu	Arg	Val	Pro	Glu	Thr	Arg	Gly	Arg	Thr	Phe	Asp	Gln	Ile	Ser
465					470					475					480
Ala	Ala	Phe	His	Arg	Thr	Pro	Ser	Leu	Leu	Glu	Gln	Glu	Val	Lys	Pro
			485					490						495	
Ser	Thr	Glu	Leu	Glu	Tyr	Leu	Gly	Pro	Asp	Glu	Asn	Asp			
			500					505							

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 500 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Met	Glu	Gln	Gln	Asp	Gln	Ser	Met	Lys	Glu	Gly	Arg	Leu	Thr	Leu	Val
1				5					10					15	
Leu	Ala	Leu	Ala	Thr	Leu	Ile	Ala	Ala	Phe	Gly	Ser	Ser	Phe	Gln	Tyr
			20					25					30		
Gly	Tyr	Asn	Val	Ala	Ala	Val	Asn	Ser	Pro	Ala	Leu	Leu	Met	Gln	Gln
		35					40					45			
Phe	Tyr	Asn	Glu	Thr	Tyr	Tyr	Gly	Arg	Thr	Gly	Glu	Phe	Met	Glu	Asp
	50				55					60					
Phe	Pro	Leu	Thr	Leu	Leu	Trp	Ser	Val	Thr	Val	Ser	Met	Phe	Pro	Phe
	65				70					75					80
Gly	Gly	Phe	Ile	Gly	Ser	Leu	Leu	Val	Gly	Pro	Leu	Val	Asn	Lys	Phe
			85					90					95		
Gly	Arg	Lys	Gly	Ala	Leu	Leu	Phe	Asn	Asn	Ile	Phe	Ser	Ile	Val	Pro
			100					105					110		

Ala	Ile	Leu	Met	Gly	Cys	Ser	Arg	Val	Ala	Thr	Ser	Phe	Glu	Leu	Ile
		115					120					125			
Ile	Ile	Ser	Arg	Leu	Leu	Val	Gly	Ile	Cys	Ala	Gly	Val	Ser	Ser	Asn
		130					135				140				
Val	Val	Pro	Met	Tyr	Leu	Gly	Glu	Leu	Ala	Pro	Lys	Asn	Leu	Arg	Gly
		145				150				155					160
Ala	Leu	Gly	Val	Val	Pro	Gln	Leu	Phe	Ile	Thr	Val	Gly	Ile	Leu	Val
			165						170					175	
Ala	Gln	Ile	Phe	Gly	Leu	Arg	Asn	Leu	Leu	Ala	Asn	Val	Asp	Gly	Trp
			180					185					190		
Pro	Ile	Leu	Leu	Gly	Leu	Thr	Gly	Val	Pro	Ala	Ala	Leu	Gln	Leu	Leu
		195					200					205			
Leu	Leu	Pro	Phe	Phe	Pro	Glu	Ser	Pro	Arg	Tyr	Leu	Leu	Ile	Gln	Lys
		210					215				220				
Lys	Asp	Glu	Ala	Ala	Ala	Lys	Lys	Ala	Leu	Gln	Thr	Leu	Arg	Gly	Trp
		225				230				235					240
Asp	Ser	Val	Asp	Arg	Glu	Val	Ala	Glu	Ile	Arg	Gln	Glu	Asp	Glu	Ala
			245					250					255		
Glu	Lys	Ala	Ala	Gly	Phe	Ile	Ser	Val	Leu	Lys	Leu	Phe	Arg	Met	Arg
		260						265					270		
Ser	Leu	Arg	Trp	Gln	Leu	Leu	Ser	Ile	Ile	Val	Leu	Met	Gly	Gly	Gln
		275					280					285			
Gln	Leu	Ser	Gly	Val	Asn	Ala	Ile	Tyr	Tyr	Tyr	Ala	Asp	Gln	Ile	Tyr
		290				295					300				
Leu	Ser	Ala	Gly	Val	Pro	Glu	Glu	His	Val	Gln	Tyr	Val	Thr	Ala	Gly
		305				310				315					320
Thr	Gly	Ala	Val	Asn	Val	Val	Met	Thr	Phe	Cys	Ala	Val	Phe	Val	Val
			325						330				335		
Glu	Leu	Leu	Gly	Arg	Arg	Leu	Leu	Leu	Leu	Leu	Gly	Phe	Ser	Ile	Cys
			340					345					350		
Leu	Ile	Ala	Cys	Cys	Val	Leu	Thr	Ala	Ala	Leu	Ala	Leu	Gln	Asp	Thr
		355					360					365			
Val	Ser	Trp	Met	Pro	Tyr	Ile	Ser	Ile	Val	Cys	Val	Ile	Ser	Tyr	Val
		370				375				380					
Ile	Gly	His	Ala	Leu	Gly	Pro	Ser	Pro	Ile	Pro	Ala	Leu	Leu	Ile	Thr
		385				390				395					400
Ile	Phe	Leu	Gln	Ser	Ser	Arg	Pro	Ser	Ala	Phe	Met	Val	Gly	Gly	Ser
			405						410				415		
Val	His	Trp	Leu	Ser	Asn	Phe	Thr	Val	Gly	Leu	Ile	Phe	Pro	Phe	Ile
		420						425					430		
Gln	Glu	Gly	Leu	Gly	Pro	Tyr	Ser	Phe	Ile	Val	Phe	Ala	Val	Ile	Cys
		435					440					445			
Leu	Ile	Thr	Thr	Ile	Tyr	Ile	Phe	Leu	Ile	Val	Pro	Glu	Thr	Lys	Ala
		450				455				460					
Lys	Thr	Phe	Ile	Glu	Ile	Asn	Gln	Ile	Phe	Thr	Lys	Met	Asn	Lys	Val
		465				470				475					480
Ser	Glu	Val	Tyr	Pro	Glu	Lys	Glu	Glu	Leu	Lys	Glu	Leu	Pro	Pro	Val
			485						490				495		
Thr	Ser	Glu	Gln												
			500												

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

TGTTTCCTAG TCTTTGCTAC A

21

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: primer

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

TTGTTAAGGC CTTCCATT

18

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 493 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Xaa	Xaa	Gly	Phe	Gln	Xaa	Gly	Ser	Val	Thr	Gly	Thr	Leu	Val	Leu
1				5					10					15	
Ala	Val	Leu	Ile	Ala	Ala	Leu	Gly	Ser	Phe	Gln	Tyr	Gly	Tyr	Asn	Leu
			20					25					30		
Gly	Val	Ile	Asn	Ala	Pro	Gln	Lys	Val	Ile	Glu	Ala	Phe	Tyr	Glu	Thr
		35					40					45			
Trp	Leu	Gly	Arg	Xaa	Gly	Glu	Xaa	Pro	Ser	Val	Pro	Thr	Leu	Thr	Leu
	50					55					60				
Leu	Trp	Ser	Leu	Ser	Val	Ser	Ile	Phe	Ala	Val	Gly	Gly	Met	Ile	Gly
	65				70					75				80	
Ser	Phe	Leu	Val	Gly	Xaa	Ile	Gly	Asn	Arg	Leu	Gly	Arg	Lys	Xaa	Ala
			85					90					95		
Met	Leu	Val	Asn	Asn	Val	Leu	Ala	Ile	Ala	Gly	Gly	Leu	Leu	Met	Gly
			100				105						110		
Leu	Ala	Lys	Xaa	Ala	Xaa	Ser	Phe	Glu	Met	Leu	Ile	Leu	Gly	Arg	Phe
		115					120					125			
Ile	Ile	Gly	Leu	Tyr	Cys	Gly	Leu	Ser	Ser	Gly	Val	Val	Pro	Met	Tyr
	130					135					140				
Val	Gly	Glu	Ile	Ser	Pro	Thr	Ala	Leu	Arg	Gly	Ala	Leu	Gly	Thr	Leu
	145				150					155				160	
Asn	Gln	Leu	Gly	Ile	Val	Ile	Gly	Ile	Leu	Ile	Ala	Gln	Val	Leu	Gly
			165					170					175		
Leu	Asp	Ser	Leu	Leu	Gly	Asn	Glu	Ser	Leu	Trp	Pro	Leu	Leu	Leu	Gly
			180				185					190			
Leu	Thr	Gly	Val	Pro	Ala	Leu	Leu	Gln	Leu	Leu	Leu	Leu	Pro	Phe	Cys
		195				200						205			
Pro	Glu	Ser	Pro	Arg	Tyr	Leu	Ile	Asn	Lys	Asn	Glu	Glu	Ala	Arg	
	210					215				220					
Ala	Lys	Lys	Ala	Leu	Gln	Arg	Leu	Arg	Gly	Thr	Ala	Asp	Val	Ser	Gln
	225				230					235				240	
Glu	Val	Ala	Glu	Met	Lys	Asp	Glu	Ser	Arg	Xaa	Met	Xaa	Ser	Glu	Lys
			245					250						255	

00001947-101601

Xaa	Val	Ser	Val	Leu	Glu	Leu	Ph	Arg	Ser	Arg	Xaa	Tyr	Arg	Gln	Pro
			260					265					270		
Val	Ile	Ile	Ala	Ile	Val	Leu	Gln	Leu	Ser	Gln	Gln	Leu	Ser	Gly	Ile
			275				280					285			
Asn	Ala	Val	Phe	Tyr	Tyr	Ser	Thr	Ser	Ile	Phe	Glu	Lys	Ala	Gly	Val
	290					295					300				
Gly	Gln	Pro	Val	Tyr	Ala	Thr	Ile	Gly	Ala	Gly	Val	Val	Asn	Thr	Val
	305				310					315				320	
Phe	Thr	Val	Val	Ser	Val	Phe	Val	Val	Glu	Arg	Ala	Gly	Arg	Arg	Thr
				325					330				335		
Leu	His	Leu	Leu	Gly	Leu	Gly	Gly	Met	Ala	Gly	Cys	Ala	Val	Leu	Met
			340					345					350		
Thr	Ile	Ala	Leu	Ala	Leu	Leu	Asp	Gln	Val	Pro	Trp	Met	Ser	Tyr	Val
		355					360					365			
Ser	Ile	Val	Ala	Ile	Phe	Gly	Phe	Val	Ala	Phe	Phe	Glu	Val	Gly	Pro
	370					375					380				
Gly	Pro	Ile	Pro	Trp	Phe	Ile	Val	Ala	Glu	Leu	Phe	Ser	Gln	Gly	Pro
	385				390					395				400	
Arg	Pro	Ala	Ala	Ile	Ala	Val	Ala	Gly	Phe	Ser	Asn	Trp	Thr	Ser	Asn
				405					410					415	
Phe	Ile	Val	Gly	Leu	Leu	Phe	Gln	Tyr	Ile	Ala	Glu	Leu	Leu	Gly	Pro
			420					425						430	
Tyr	Val	Phe	Ile	Val	Phe	Ala	Val	Leu	Leu	Leu	Leu	Phe	Phe	Ile	Phe
		435					440					445			
Thr	Phe	Leu	Lys	Val	Pro	Glu	Thr	Lys	Gly	Arg	Thr	Phe	Asp	Glu	Ile
	450					455					460				
Ala	Ala	Ala	Phe	Arg	Lys	Xaa	Asn	Lys	Xaa	Glu	Gln	Pro	Glu	Lys	Glu
	465				470					475				480	
Ser	Ile	Glu	Glu	Leu	Glu	Pro	Leu	Gly	Pro	Asp	Glu	Xaa			
				485					490						